

University of Groningen

Distribution network monitoring

Blaauwbroek, Niels; Kuiken, Dirk; H. Nguyen, Phuong; Vedder, Hans; Roggenkamp, Martha; Slootweg, Han

Published in:
Energy Policy

DOI:
[10.1016/j.enpol.2017.12.041](https://doi.org/10.1016/j.enpol.2017.12.041)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Blaauwbroek, N., Kuiken, D., H. Nguyen, P., Vedder, H., Roggenkamp, M., & Slootweg, H. (2018). Distribution network monitoring: Interaction between EU legal conditions and state estimation accuracy. *Energy Policy*, 115, 78–87. <https://doi.org/10.1016/j.enpol.2017.12.041>

Copyright

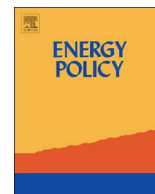
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Distribution network monitoring: Interaction between EU legal conditions and state estimation accuracy

Niels Blaauwbroek^{a,*}, Dirk Kuiken^b, Phuong Nguyen^a, Hans Vedder^b, Martha Roggenkamp^b, Han Slootweg^a

^a Eindhoven University of Technology, Eindhoven, the Netherlands

^b University of Groningen, Groningen, the Netherlands

ARTICLE INFO

Keywords:

Electrical distribution network
Measurement data
Legal framework
Privacy
State estimation
Monitoring accuracy

ABSTRACT

The expected increase in uncertainty regarding energy consumption and production from intermittent distributed energy resources calls for advanced network control capabilities and (household) customer flexibility in the distribution network. Depending on the control applications deployed, grid monitoring capabilities that accurately capture the system operation state are required. In order to establish such monitoring capabilities, several technical and legal challenges relating to monitoring accuracy, user privacy, and cost efficiency need to be tackled. As these aspects have complex mutual interdependencies, a universal approach for realising distribution network monitoring is not straightforward. Therefore, this article highlights these issues and proposes a method to evaluate monitoring accuracy and the proportionality of personal data processing, and to illustrate the interdependencies between finding the legal grounds for data processing and the monitoring accuracy the processed data produces. To illustrate the method, several test cases are presented, in which the accuracy of network monitoring is assessed for different measurement configurations, followed by an analysis on the legality of the configurations.

1. Introduction

The energy landscape is changing. Renewables are being integrated at an ever-increasing pace into our energy systems, and much of the current energy demand is electrifying due to e.g. heat pumps and electric vehicles. Because of the intermittent character of electricity produced from renewable energy sources (RES), the production of electricity becomes less predictable and controllable. Both the electrification and the integration of RES increase the peaks that exist on the electricity networks. As current electricity network capacity is based on the maximum power peak, an increase in peak power is always translated into additional investments in the current electricity networks' capacity. These additional investments are expected to be extremely significant (Verbong et al., 2016).

A more cost-efficient solution for dealing with the increasing uncertainties on the distribution networks would be to utilize the current electricity system in a more flexible way. This could be realized by developing advanced network operation strategies, i.e. active distribution networks, to keep the network within efficient, stable and safe operation conditions. Examples of such strategies are power flow control algorithms for optimising the network states using local controllers

for reactive power and voltage, or demand side management programs and market mechanisms for using flexibility from both active power production and consumption (Blaauwbroek and Nguyen, 2015; Gungor et al., 2013; Torbaghan et al., 2016). However, in order to realise these future network operation strategies, advanced monitoring capabilities for the distribution network to accurately capture its actual system states are required (Angioni et al., 2016; Pérez-Arriaga, 2013). The information gained from network monitoring will serve as input for various network operation strategies to operate the network more efficiently and within secure boundaries. However, these monitoring applications will rely on various data sources, such as network measurements, pseudo-measurements, weather forecasts and end user data. The accuracy requirements for this data depend on the goals and functionalities the advanced network operation strategies are supposed to realise. In any case, data collection is costly, as investments have to be made in monitoring equipment, communication infrastructure, etc. Therefore, hardly any measurement equipment is installed in the current distribution network, preventing Distribution System Operators (DSOs) from gaining insight in their system states. Usually, the investments that have to be made for gaining this insight can be related to the degree of accuracy (quality): the higher the accuracy (e.g. shorter

* Corresponding author.

E-mail addresses: n.blaauwbroek@tue.nl (N. Blaauwbroek), d.kuiken@rug.nl (D. Kuiken).

timeframe of measurements, more detail, and reliability), the higher the investments costs (Singh et al., 2009). In addition, currently a high level of legal ambiguity with regard to data processing in distribution networks exists, which is a hurdle for realising network monitoring (EDSO for Smart Grids, 2015; European Energy Regulators, 2015). Most of this ambiguity can be ascribed to lack of a clear framework for assessing the legality of data processing for network monitoring purposes. Also the lack of clear and measurable goals makes it difficult to assess whether, how, and which data should be processed. Without a clear and justifiable framework for data processing, it is impossible to process personal data for network monitoring purposes. Therefore, next to making investments in measuring equipment, the legal conditions regarding the measuring, processing, and estimation of information in relation to system operation need to be clarified. These legal conditions mainly relate to two aspects.

Firstly, European Union (EU) law requires that DSOs provide for secure, reliable and (cost-)efficient electricity networks. In practice this means that within their framework of (legal) requirements, DSOs should strive for optimal efficiency of their electricity networks. In this context network monitoring is also subject to the requirements of keeping networks secure and reliable in a (cost-)efficient manner. Consequently, the costs of monitoring and control applications used for network operation should be proportionate (cost-efficient) in relation to the benefits (security, reliability or efficiency) they create.

Secondly, while monitoring their network, DSOs have to respect the privacy of their (household) customers as much as possible, especially taking into account that household customers generally become more vulnerable to (unlawful) privacy breaches if the network is equipped with advanced monitoring capabilities (Milaj and Mifsud Bonnici, 2016). Although network monitoring might contribute to more secure, reliable and efficient networks, they might also reduce household customer privacy. Therefore, a balance has to be struck between both interests.

Considering both the technical and legal aspects involved in network monitoring as discussed above, it is clear that the complex interactions amongst these aspects complicate the question on how the monitoring functionality can be realized for a specific case. Therefore, the aim of this article is to discuss these issues and introduce a method that is able to strike a balance between the interests of the DSO and household customers, resulting in a legally feasible outcome with the lowest monitoring error margin (highest data quality) and reasonable costs for installing measurement equipment.

The article is structured as follows. To begin with, Section 2 introduces the technical aspects and goals of the distribution system monitoring and the functions it should serve. Section 3 provides the legal framework for data capturing in distribution networks (including a short introduction to the newly adopted EU General Data Protection Regulation – GDPR). Section 4 introduces the method to evaluate monitoring accuracy and the proportionality of processing personal data. Section 5 discusses a number of test-cases, for which the performance of monitoring applications is assessed and the legal feasibility of the test-cases is analysed. Finally, the article concludes in Section 6.

2. Technical aspects in distribution system monitoring

As aforementioned, due to the current lack of monitoring capabilities in distribution networks, newly developed monitoring applications are required in order to establish adequate control capabilities in distribution networks. These monitoring applications will give insight in the system states of the network. The system states form a data set that defines the operation state of the network uniquely. In order to acquire the system states, they can be measured directly (e.g. voltage magnitude levels and phases) at each node of the grid with high measurement frequencies. However, the installation of measurement and communication equipment for the large number of nodes in distribution networks will be a costly exercise (especially for phasor measurements).

Besides, this measurement data might not always be as accurate as required (e.g. because of failing measurement equipment or communication delays/losses). Therefore, as commonly applied in transmission systems, state estimation of distribution systems has been proposed (Della Giustina et al., 2014) in order to enhance the accuracy and reliability of the monitoring in distribution systems. The next paragraphs further explain the background of state estimation, network observability and the possible types of measurement equipment and data.

2.1. Power system state estimation

State estimation is a process to obtain the maximum likelihood estimate of the system states, based on measurements, pseudo-measurements (e.g. historical data from other sources) and a model of the network. The model of the network and the system states together uniquely define the full operation state of the network. Estimation of the system states is applied because usually not all the system states can be measured directly, or the measurements values might be inaccurate. Instead of relying on inaccurate measurements or pseudo-measurements, a more accurate estimate of the true system states can be obtained by using a model of the network and a state estimation algorithm. This way, less measurements are required or a less expensive metering infrastructure can be used. Algorithms for bad data detection can identify faulted sensors or data that is arriving late, such that the data can be excluded from the state estimation process and eventually be replaced by pseudo-measurements to retain observability of the network. For a fully observable network, typically the system states are defined as the set of all nodal voltages and corresponding angles, but also the set of all branch currents and angles can be used, together with a reference voltage. The last option is gaining more attention in recent research, because of its computational performance (Pau et al., 2013; Wang and Schulz, 2004). Before state estimation can be carried out, first the observability criteria of the network need to be satisfied.

2.2. Network observability

In order to make the network fully observable, a minimum number of measurements is required. This minimum number is related to the number of nodes or branches in the network. A branch is formed by a line or cable of the network, whereas a node is defined as a point where multiple branches come together. Suppose the radial network consists of n nodes and therefore $b = n - 1$ branches. The number of system states that now uniquely defines the network is $2n - 1$. In case of nodal voltage state estimation, this number is made up by n voltage amplitudes of all the nodes and $n - 1$ voltage angles of all the nodes except the reference node (slack node). In case of branch current state estimation, this number is made up by b current amplitudes of all the branches, b current angles of all the branches and 1 reference voltage amplitude of the slack node. Now, for a fully observable network, at least $2n - 1$ measurements are required that can be mapped independently to the $2n - 1$ system states. Therefore, if insufficient measurement equipment is installed to provide this data, the measurements have to be complemented with pseudo-measurements. In the test cases presented in Section 5, the data set that is input to the state estimation at least contains $2n - 1$ measurements and pseudo-measurements.

2.3. Network measurements

In practice, the (pseudo) measurements themselves can be obtained from various sources. These sources can include data from measurement equipment installed in the distribution network, as well as all kinds of pseudo information in case real measurement data is missing (because of a lack of measurement equipment, bad data connections etc.).

Real measurement data can be obtained from measurement

equipment installed by the DSO, or from measurement equipment at the customer side. Mostly, measurement equipment installed by the DSO will be located in the substation and other important junctions of the network. Measurement equipment at the customer side can include smart meters or eventually measurement components that are part of controllers such as inverters for photo voltaics (PV) installations or batteries of electric vehicles (EV).

Pseudo measurements can be derived from various data sources, such as historical power consumption data, weather forecasts, household/customer details (number of inhabitants, available appliances) etc. The key importance is that this data can be used to reconstruct power injection profiles (consumption or production) at the customer connection point, where real measurement data regarding power injection is missing. Although this information is likely to be highly inaccurate, it still helps to restore network observability. In order to compensate for the relatively high variance of the pseudo measurements, in the state estimation algorithm the pseudo measurements are taken into account with a lower weight compared to real measurements as specified in [Section 4](#).

In any case, regardless of whether data is collected directly from the network or its users, or indirectly from other sources, legal conditions apply to the processing of such data. These conditions have to be assessed in order to define if system operators are expected or required to process data in the first place, and if so, which data can be collected based on these expectations or requirements. As such, the following section continues with discussing the applicable legal framework for network monitoring in distribution systems.

3. Legal framework for distribution network monitoring

In the EU, all Member States have to ensure that every system user (consumer or producer) is entitled to use the electricity at non-discriminatory conditions according to article 32 of the Electricity Directive ([European Union, 2009](#)). Moreover, household consumers (representing the majority of system users) have a right to enjoy a universal service, “that is the right to be supplied with electricity of a specified quality within their territory at reasonable, easily and clearly comparable, transparent and non-discriminatory prices.”, which also implies the right to use the electricity system (article 3(3) Electricity Directive). In order to ensure the right of all household customers to use the electricity system, DSOs have the general task to ensure secure, reliable and efficient electricity systems (article 25 Electricity Directive). Both rights have to be implemented into national legislation.

Although currently DSOs are generally not required (from a legal point of view) to produce detailed measurements in their distribution networks, this situation is expected to change. Given the high amount of uncertainty expected in future distribution networks, a minimum level of network monitoring seems crucial for ensuring future ‘secure, reliable and efficient’ network operation ([Diestelmeier and Kuiken, 2017](#)). Nevertheless, such monitoring requires data to be collected, most likely data from or about household customers.

In addition, in the EU, everyone has the right for their private and family life to be respected and its personal data to be protected ([Council of Europe, n.d.](#); [European Union, 2012a, 2012b](#)). In order to ensure this right, the General Data Protection Regulation (GDPR) ([European Union, 2016](#)) lays down specific rules on the protection of data (article 1 and recital 1 of the GDPR). From 25 May 2018 onwards, the GDPR will be directly applicable in all EU Member States (articles 99(2) and 94(1) GDPR), and will replace the current Data Protection Directive. The GDPR provides the exact same legal framework for all EU Member States.

In order to analyse how the GDPR works and how it relates to network monitoring, we first define the roles and definitions applicable in the GDPR in relation to the electricity sector. Thereafter, we define what would classify as personal data, followed by an analysis of the (pre)conditions for data processing.

3.1. Definitions and roles

In order for the GDPR to apply, personal data must be at stake. Personal data is data that relates to either an identified or identifiable person, which can also be a household (article 4(1) GDPR, and ([Article 29 Data Protection Working Party, 2007](#)). Such a person (household) is referred to as the data subject (the subject), whilst the entity processing the data is referred to as the data processor (the processor). The processor ‘processes’ the data, meaning for example the collecting, storing, organizing, altering, disseminating, or erasing of data (article 4(2) GDPR). Next to the data subject and the data processor, a data controller (the controller) can be identified. The controller sets the goals and means used for the processing of data (article 4(7) GDPR). For the scope of this article, the data processor and controller are further referred to as the DSO, and to the data subject as household customer.

3.2. Classification of ‘personal data’

As a general rule, personal data should always relate to individuals or identifiable persons. However, it is not required that personal data is collected from the household customer itself. Data sets (e.g. profiles) can also be based on non-personal data, such as pseudo-data. For example, (smart) meter data from subject A can be processed, pseudonymised, and translated into a general profile fitting to the personal conditions of subject A. In making such a profile many information sources could be used, both public and non-public, such as data from the network connection register. In principle, the profile of subject A can also be applied to other individuals that are in a similar or comparable position as subject A. Such profiles could perhaps even be produced within a fairly accurate range of measured data. If such profiles are attributed (e.g. based on name, EAN code, location data, etc. Article 4(1) GDPR) to household customers, such profiles can be considered to be personal data.

3.3. (Pre)conditions for processing personal data

The main purpose of the GDPR is to balance the interests of the data processor/controller and the data subject. In our particular case, a balance should be struck between the interests of the DSO and the interests of the household customer. Clearly, the interest of household customers is to have their privacy protected as much as possible. In turn, the DSO only process personal data if having a (legal) ground (legitimate interest) for processing. Consequently, from the DSOs perspective the first requirement of the GDPR is to establish a legitimate ground for processing personal data.

A legitimate ground can either be based on the consent of the subject (article 7 GDPR), or formed by the requirements for the performance of a contract, a legal obligation, the protection of vital interests of the data subject, a task of public interest or the official authority of the controller, or a legitimate interest of the controller (article 6(1) GDPR). When zooming in on the possible grounds for DSOs to process personal data for creating network observability, both the tasks of the DSOs, as well as the right for household customers to be provided with system services need to be considered. When considering the potential grounds, the first and clearest ground is based on consent by the household customer. In this particular case an explicit agreement between the DSO and the household customer exists to process personal data. However, even without such an explicit ‘agreement’, an implicit agreement can exist if processing personal data would be necessary for the DSO to perform its above described tasks, or to perform its legal obligations based on a contract with a household customer to provide that user with system services. Arguably, processing could also be based on either the vital interest of the household customer (to be supplied with affordable system services), or the public interest (affordability of the electricity system for all users). Finally, processing could also be based on a legal ground in national law (e.g. as a result of implementing

the Electricity Directive), requiring of prescribing the DSO to process personal data for enhancing network observability. However, it is important to note that if processing of personal data would be allowed, DSOs are not provided with a *carte blanche*; they will always have to fulfil the conditions by the GDPR for data processing and minimize their processing activities as much as possible (Article 29 Data Protection Working Party, 2011).

After the ground for processing has been established, DSOs must most likely perform an impact assessment before starting the processing (article 35 GDPR). The assessment must include a description of how data is going to be processed, the necessity and proportionality of the processing in relation to the purpose of processing, and the data protection risks that are associated with the processing (Bieker et al., 2016). In order to increase legal certainty for the DSOs, they can draft a code of conduct, to be approved by either the (national) supervisory authority, or by the European Commission if the code would relate to data processing in multiple Member States (article 40 GDPR).

Once processing personal data, DSOs have to keep record of their processing activities in the form of a processing administration (article 30 GDPR, and see (de Hert and Papakonstantinou, 2016)). In turn, household customers have a right to know whether (their) data is being processed, to view and review such data, and to object to both the content of processed data and the processing itself (Chapter III GDPR), in order to being able to protect their own interests.

3.4. Minimization of personal data

As a general requirement, processing personal data should not go beyond the purpose for which it is collected (article 5 GDPR). As such, the processing of personal data must be minimized as much as possible. One way of minimizing the processing of personal data is for example by decreasing the time intervals in which data is collected; the higher the frequency, the more difficult it becomes to justify the processing of such data (McKenna et al., 2012).

Nevertheless, if personal data is collected, DSOs should also consider to ‘pseudonymise’ data. If possible, the DSO should pseudonymise the data as much and as soon as possible (recital 78 and 156 GDPR). Pseudonymising data means that the data cannot (directly) be traced back to the subject without additional information (e.g. by separating meter measurements and corresponding EAN codes). Pseudonymisation can be performed by applying technical (e.g. encryption) or organisational measures to disconnect the data and the subject (article 4(5) GDPR).

Once personal data is pseudonymised, such data is considered to be non-personal. As such, the processing of pseudonymised data does not fall in the scope of the GDPR. However, it should be taken into consideration that pseudonymised data can also be reversed to personal data. If data can be (re)attributed to a specific data subject, for example by combining the data with other information, such data is still considered to be personal data (Council of Europe, 2014).

4. Method

The method applied to evaluate the interdependencies of monitoring accuracy and the legality of processing of personal data, consists of several steps detailed in this section, consisting of two main steps. The first main step is analysing the technical performance of a specified monitoring configuration. The second main step is assessing the legal feasibility of a specified monitoring configuration. In this, it is argued that, because the monitoring accuracy depends on the availability of real measurements – and therefore customer data – the monitoring accuracy is influenced by the legal conditions regarding processing of personal data. However, as EU law allows the processing of personal data only if it is proportional for the goal of processing the data (i.e. for example reaching a minimum level of monitoring accuracy for specific control applications), the monitoring applications also influences the

conditions for processing data. These interactions are discussed in more detail in the following subsections.

4.1. Step 1: assessment of monitoring accuracy

For the performance evaluation of state estimation algorithms for different cases, power flow calculations of a test network are used as a reference of the system state of the network. This way, the true system states of the network can be obtained, which can serve as a reference to which the outputs of the state estimation can be compared. From these true system states, measurements are taken according to different measurement models divided over different cases. Together with pseudo-measurements from various sources, they form the input for the state estimation. The estimated system states resulting from the state estimation algorithm are compared with the true system states from the power flow calculations, after which the deviation (error) between the true and estimated system states is determined.

4.1.1. Network model

The network model used in these simulations is the IEEE European LV test feeder (IEEE Power and Energy Society, 2015). This is a three-phase distribution network counting 117 nodes and 116 branches with specified line parameters and 55 households tapping of the feeders, as displayed in Fig. 1. The model has been implemented in MATLAB/Simulink, allowing to run a time horizon simulation of the distribution network. For the time horizon simulations as well as computing the pseudo measurement variances, many 24-h load profiles are required from several years for each of the households. As the publication of the IEEE European LV test feeder is lacking this data, similar data has been obtained from real measurement data as published by the Pecan Street project (Pecan Street Inc., 2016). From the data available in the Pecan Street project, a group of 55 households is selected of which their load profiles have been measured at the 21st of June in the years from 2012 to 2016. This selection reflects a representable group of households with various types of loads, including significant amounts of PV and EV. For each test case discussed in Section 5, the same profiles have been applied for each household. Differences exist however in the measurement configuration and usage of pseudo-measurements for each of the test cases.

4.1.2. Measurement models

Overall, we distinguish three types of measurements. Firstly, measurements obtained from equipment installed by the DSO is concerned. This measurement equipment is assumed to have high accuracy and

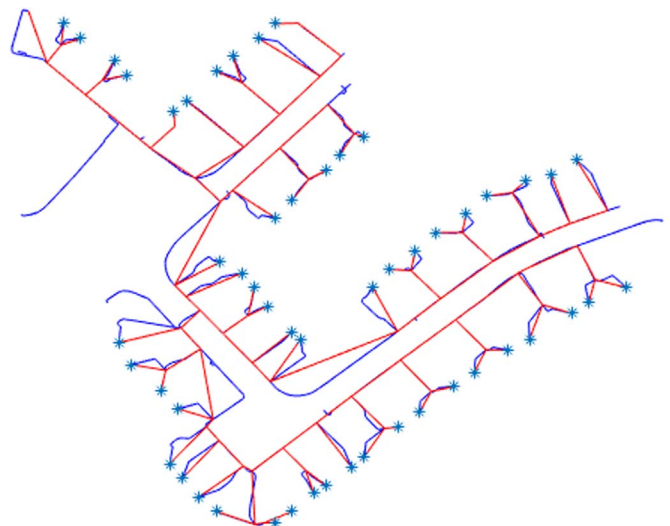


Fig. 1. IEEE European 55-node low voltage distribution feeder.

reliability with a 1-min measurement interval.

The second category concerns measurements obtained from smart meters installed in the customer households. These measurements have various measurement intervals, where the measurement intervals can be either synchronised or unsynchronised. In the case of synchronised measurement intervals, the smart meters are assumed to all take measurements at the same moment. For example, in case of a 15-min measurement interval, measurements could be taken at 0, 15, 30 and 45 min past the whole hour. For the unsynchronised case, the smart meters take their measurements at a random moment within the time interval, where the assumption is that the number of smart meters are equally distributed over the time interval. Independent of either of the two methods, within the simulation the measurement values are derived during runtime from the true system states as simulated by MATLAB/Simulink. First, the measurement (e.g. power injection) value is calculated out of the simulated system states. After this, a zero-mean, normally distributed random error is added to this measurement value to model the inaccuracies of the measurement equipment, before inputting it to the state estimation algorithm. The errors applied to the measurements are considered uncorrelated.

The last category of measurements concerns pseudo-measurements. These are measurements that can be substituted in the absence of real measurement data. Pseudo-measurements are compiled offline, before running the simulation. They consist of historical measurement data for the specific day of the year, averaged over 5 years and all households in the network for each 1-min time interval.

4.1.3. Weighted least squares state estimation algorithm

The state estimation algorithm used in this paper is a branch current state estimation algorithm based on the work presented in Wang and Schulz (2004). It is based on the weighted least squares method (WLS) that maximizes the conditional probability function of the system states. As such, WLS state estimation is solved by minimizing

$$[z - h(x)]^T R^{-1} [z - h(x)] \quad (1)$$

In this equation, x is the state vector containing the system states uniquely defining the operation state of the network and h is the function relating the state vector x with the measurement vector z according to

$$z = h(x) + e \quad (2)$$

Finally, R is the covariance matrix of the normally distributed measurement error e . In practice, the WLS estimate of (1) is obtained using the Newton-Raphson method by iteratively solving (1)

$$x^{k+1} = x^k - G(x^k)^{-1} g(x^k) \quad (3)$$

where

$$g(x) = \frac{\partial J(x)}{\partial x} = H^T(x) R^{-1} [z - h(x)] \quad (4)$$

and the gain matrix $G(x^k)$ is the Jacobian of $g(x^k)$ evaluated around x^k :

$$G(x^k) = \frac{\partial g(x^k)}{\partial x} = H^T(x^k) R^{-1} H(x^k) \quad (5)$$

The measurement Jacobian matrix $H^T(x^k)$ is obtained by differentiating the measurement function $h(x)$ at x^k , of which many examples for different power system models and system states can be found in literature.

4.1.4. (Pseudo) measurement variances

In order to perform a proper state estimation for the distribution network, the variances of the input measurements need to be known. For the measurement equipment as installed by the DSO, the variance will be a property of the measurement equipment itself. For the measurements used in this paper, the variance is assumed to be 10^{-5} . For pseudo measurements and smart meter measurements however, the

variance needs to be derived in another way.

For the power injection pseudo measurements, the variance for each 1-min time interval is based on the profiles for the different households for five different years. As detailed before, the pseudo measurement is calculated as the minute interval household consumption (negative for production) $c_{h,m}^y$ in minute m averaged over the households $h \in \mathcal{H}$ and the years $y \in \mathcal{Y}$, where \mathcal{H} is the set of all households of size $|\mathcal{H}| = H$ and \mathcal{Y} is the set of all years of size $Y = |\mathcal{Y}|$. From this, it follows that the variance for the pseudo measurement v_m for each 1-min time interval m can be calculated as:

$$v_m = \frac{1}{HY} \sum_{h \in \mathcal{H}, y \in \mathcal{Y}} \left(c_{h,m}^y - \frac{1}{HY} \sum_{h \in \mathcal{H}, y \in \mathcal{Y}} c_{h,m}^y \right)^2 \quad (6)$$

For the variance of the smart meter power injection measurements, it is important to consider that the state estimation algorithm runs on a 1-min interval base over the 24-h time span. The smart meters however are configured to have a longer measurement interval, as specified in the test cases. Obviously, at the moment the measurement is taken, the variance is given by the accuracy of the measurement equipment. However, for each moment in time between the previous measurement and the next measurement, the variance depends on the statistical change in the measurement value compared to the moment the measurement was taken. Therefore, now we calculate the variance v_k for each minute k within a measurement interval based on the difference between the power injection value $c_{h,0}^y$ at the moment the measurement is taken and the difference between the power injection value $c_{h,k}^y$ at minute k after the measurement is taken:

$$v_k = \frac{1}{HY} \sum_{h \in \mathcal{H}, y \in \mathcal{Y}} \left((c_{h,0}^y - c_{h,k}^y) - \frac{1}{HY} \sum_{h \in \mathcal{H}, y \in \mathcal{Y}} (c_{h,0}^y - c_{h,k}^y) \right)^2 \quad (7)$$

Finally, these variances are averaged over all households in the network and the five different years for each simulation time interval. More details on the numerical values can be found in Section 5.

4.1.5. Performance assessment and indicators

In order to finally compare the performance of different cases, a performance measure is needed to indicate how accurate the state estimation algorithm is compared to the true system states. The performance measure used in this paper is expressed as the absolute difference between the estimated nodal voltage phasor and the true nodal voltage phasor in percentage of the absolute estimated nodal voltage, averaged over each of the three phases and each of the nodes:

$$\text{mean}_{n \in \mathcal{N}} \left(\text{mean}_{p \in \{a,b,c\}} \left(\left| \frac{V_{\text{est}}^{n,p} - V_{\text{sim}}^{n,p}}{|V_{\text{est}}^{n,p}|} \right| * 100\% \right) \right) \quad (8)$$

4.2. Step 2: assessment of legal feasibility

In order to assess the legality of data processing, a number of requirements have to be assessed. Especially given the requirements of data protection by design and default (article 25 GDPR) and the requirement to make a data protection impact assessment (article 35 GDPR), it is crucial to assess the legality of data processing before the processing starts. For the legal feasibility, the following requirements have to be assessed. Firstly, the nature of the data to be processed is assessed. Secondly, the ground for data processing is assessed. Thirdly, the interests of the data subject in relation to the interests of the data processor are assessed.

4.2.1. Nature of data

The nature of data can be either personal or non-personal. In order to define the nature it is questioned whether data can be related to individuals or identifiable persons. If data cannot be related to

individuals or identifiable persons, such data is non-personal (see Section 3.2). Consequently, the requirements for personal data of the GDPR do not apply (European Commission, 2011).

4.2.2. Ground for processing

For this article, we assume that the (personal) data is processed in order to make state estimations as a requirement for the DSO to ensure a (more) secure, reliable and efficient electricity system in order to safeguard household customers' rights to be supplied with system services (see Section 3.3). In other words, we assume that having state estimations enables the DSO to improve or maintain the security, reliability and/or efficiency of its networks. Whether and to what extent state estimations allow for increased security, reliability and/or efficiency in practice depends on the exact case and on how the estimations are utilized.

4.2.3. Balancing of interests

In order to balance the interests of both the data subject and processor, the personal character of the data to be processed should not go beyond the legitimate interests derived from the ground for processing as described in Sections 3.3 and 3.4. As such, the personal character of processed data should be minimized, keeping in mind the question: “given the purpose of processing: how detailed should personal data be?” Answering this question requires a technical performance analysis of the processing methods in relation to the goal of processing (necessity and proportionality). In assessing the methods used in relation to the goal, also the interests of the DSO (the public interest) and household customers need to be weighed. In terms of outcome, we determine the aggregated level of processing (household level, substation level, etc.) and the frequency of processing (e.g. quarterly, minute, etc.). Also, we take into account potential alternatives: “are less intrusive (at proportionate costs) measures required and available to secure the goal?”

5. Test cases and numerical results

This section deals with the practical and numerical results for establishing monitoring in distribution systems. To this extend, various test cases are analysed. Each of the cases makes use of a different set of measurements, measurement intervals and pseudo measurements. For all cases (other than the base case), a minimum smart meter coverage of 80% is used. This minimum is based on the EU requirement for Member States of having 80% smart meter coverage in their electricity systems (Electricity Directive, Annex I(2)).

The following subsections for each case describe the measurement variances obtained from the data sets used, followed by the assessment of state estimation accuracy according to the method described in Section 4. From here, each case is analysed on monitoring accuracy and legal feasibility.

5.1. Evaluation of monitoring accuracy

5.1.1. Base case

The base case aims to mimic the monitoring accuracy for distribution networks as can be obtained with current practice for network operation, without additional investments costs for measurement equipment. The only measurement equipment installed in the network is located in the transformer substation, as is often the case nowadays. It is assumed that measurements of the substation nodal voltage, as well as the outgoing active and reactive power are available on a minute interval base. In order to establish observability, complementary pseudo-measurements in the network are taken into account in the estimation process of each household connection. Fig. 2 displays the pseudo measurement variance over time. As can be seen, this variance is quite significant, due to the high stochasticity of the load profiles applied in the test cases.

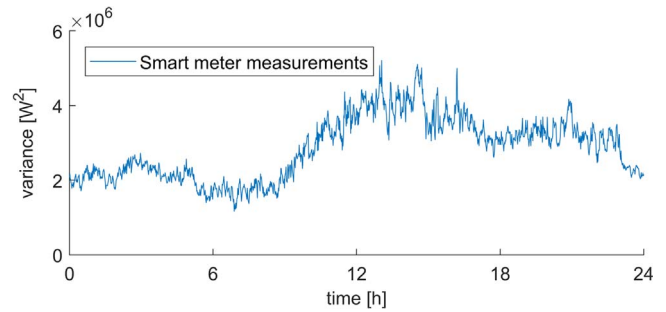


Fig. 2. Base case measurement variance.

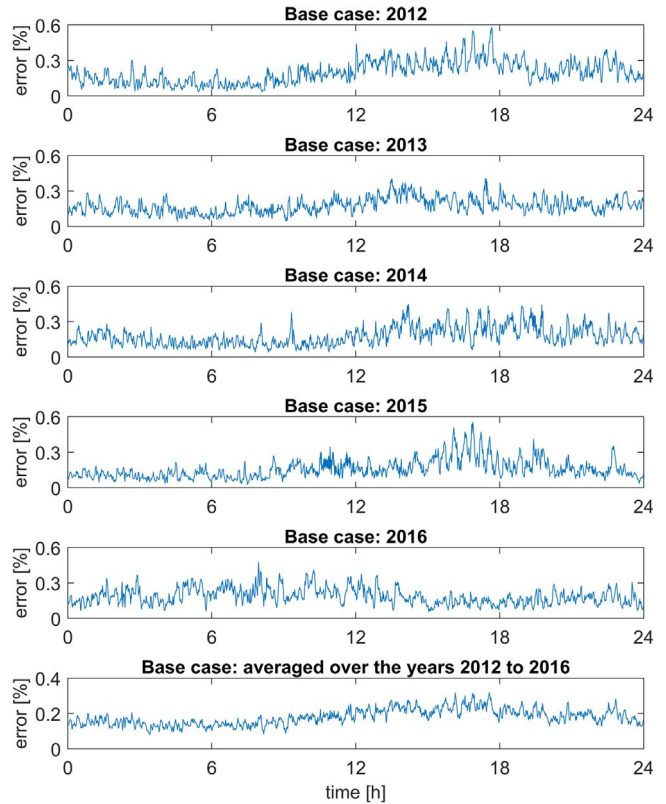


Fig. 3. Base case state estimation errors for different years.

Fig. 3 displays the mean error between the estimated voltage phasors and the true voltage phasors over a 24-h period, averaged over all nodes and phases for the years 2012–2016. The lowest graph of Fig. 3 presents the mean error, averaged over the years 2012–2016. It can be seen that the averaged error over time for the years 2012–2016 shows a strong correlation with the variance as presented in Fig. 2 and remains lower than 0,6% throughout the day. Although this might seem insignificant, the non-averaged error for individual nodes peaks around 1.8% in certain time intervals. Considering an operational range of nodal voltage in the network between 0,9 and 1,1 pu (EN 50160), this would result in an error of 9% of the control range of the network voltage. Besides, the state estimation algorithm in this work is based on ideal network models. As this will not be the case in reality, additional estimation errors might occur (Blaauwbroek et al., 2017). These two aspects together can result in a very significant error, forming a motivation for deploying a more advanced measurement configuration in the network. For this, various variants have been analysed in the next sections.

5.1.2. Variant 1

The measurements in variant 1 consist of voltage- and power

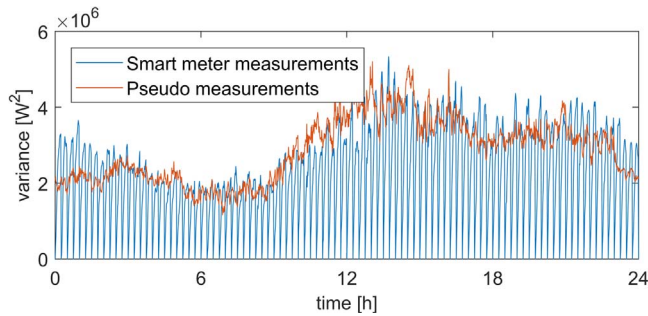


Fig. 4. 15-Minute smart meter measurement variance over time.

measurements in the substation on a minute interval base. On top of that, household power injections measured from smart meters are available on a 15-min interval base. These smart meters have a measurement accuracy with a variance of 10^{-3} times the measurement value. The smart meter measurements are complemented with pseudo-measurements to restore observability of the network. Fig. 4 displays the measurement variances for the pseudo measurements, as well as the 15-min time interval smart meter measurements. From here, it can be seen that the variance for the smart meter measurements drops to zero every 15-min as expected. However, between two smart meter measurements, the variance rapidly increases to levels comparable to those of the pseudo measurements. This is due to the high stochasticity of the individual loads in the distribution network. Consequently, we expect that smart meter measurements with 15-min intervals are inadequate for improving the monitoring accuracy in the periods between two measurements.

In order to verify this, for the variant with 15 min time interval smart meter measurements, the accuracy of three different measurement configurations are analysed. These configurations concern respectively: 1) synchronised smart meter measurements with 80% coverage (randomly distributed) as complying with the EU requirement, 2) synchronised smart meter measurements with 100% coverage, and 3) unsynchronised smart meter measurements with 100% coverage. The results of these simulations are depicted in Fig. 5, showing the errors over time between the estimated and true system states, averaged over the three phases, all nodes and the years 2012–2016. As a reference, also the averaged error of the base case (Fig. 3) is displayed in red. From the results, we can clearly see that for the first measurement configuration, the estimation error strongly reduces (but does not reach zero) each 15 min, at the moments the synchronised measurements are taken. However, in between the measurements, as expected from the measurement variances, the error is comparable to the error of the base case. This effect is even more visible for the second measurement configuration, where the estimation error approaches zero each 15 min due to the 100% coverage of smart meter measurements. For the third measurement configuration, the measurement error does not drop each 15 min due to the unsynchronised measurements. Although one might expect an improved estimation error compared to the base case when using unsynchronised smart meter measurements, due to the relatively high average variance of the smart meter readings in each minute time interval, this turns out to be not the case.

5.1.3. Variant 2

The measurements in variant 2 are similar to those of variant 1, with the only difference that the smart meter measurements are now taken with 5-min time intervals. As can be seen in Fig. 6, the variance of the 15 minute measurement intervals is lower than for the 15 minute measurement intervals and therefore lower than the variance of the pseudo measurements.

In the averaged errors displayed in Fig. 7, we see similar behaviour as in variant 1. The error drops each 5 min for the first measurement configuration with 80% coverage of synchronised smart meter

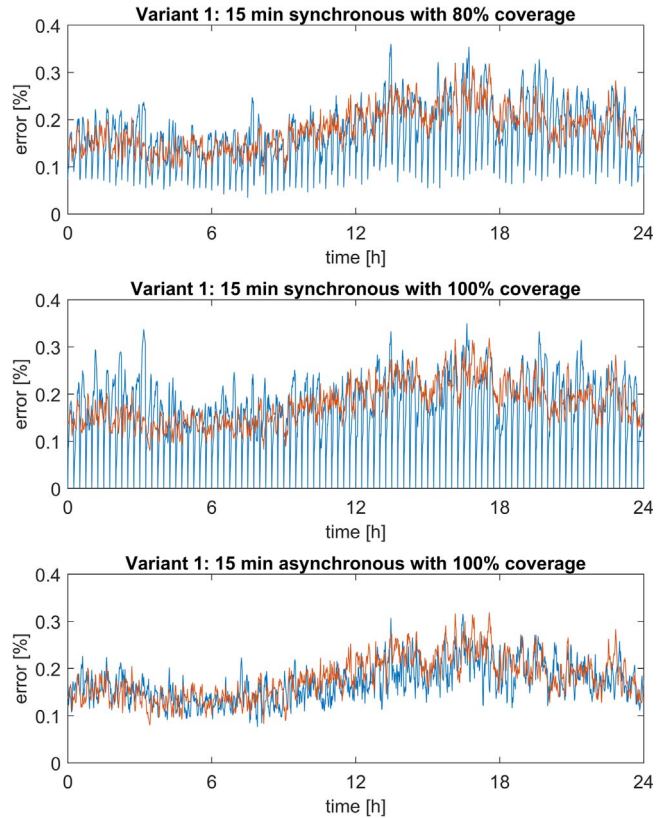


Fig. 5. Variant 1 state estimation errors for different configurations.

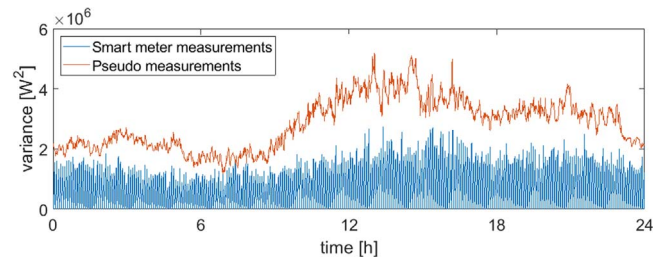


Fig. 6. 5-Minute smart meter measurement variance over time.

measurements. Similarly, for the second configuration with 100% coverage of synchronised smart meter measurements, the error approaches zero each 5 min time interval. Due to the lower variance between the measurements compared to the 15-min time interval smart meter measurements, the overall error between estimated and true system states is also lower over time. This effect is also visible in the third measurement configuration with 100% coverage of unsynchronised smart meter measurements.

5.1.4. Variant 3

In order to improve the monitoring accuracy even further, variant 3 again is similar to the previous variants, but now with a 2-min time interval for smart meter measurements. This time, the measurement variance is significantly lower as displayed in Fig. 8.

In variant 3, the trend of variant 2 continues with lower overall errors due to the lower variances of the smart meter measurements with 2-min time intervals for each of the measurement configurations, as can be clearly seen in Fig. 9.

5.1.5. Overall comparison of monitoring accuracy

Table 1 shows the full comparison of monitoring accuracy for each of the variants and all measurement configurations. The table lists

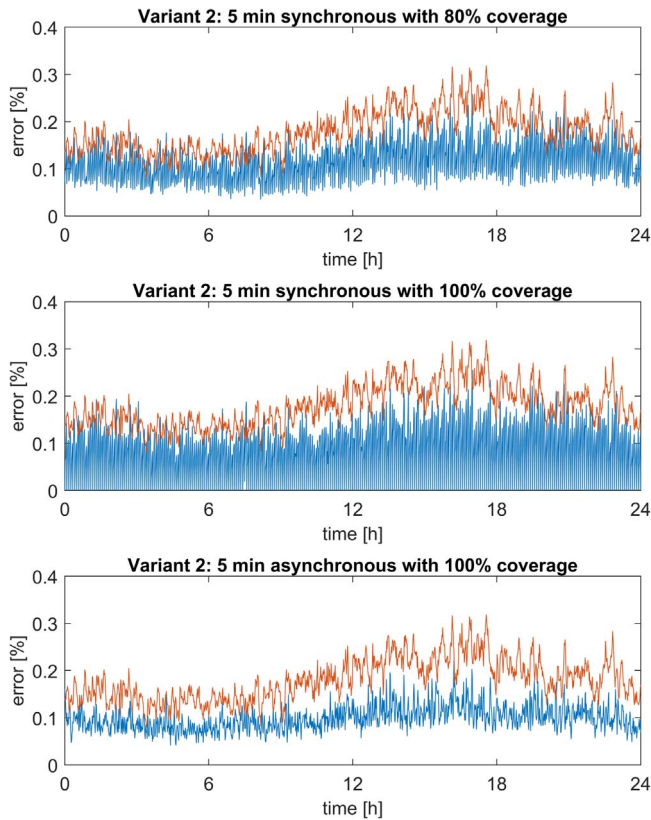


Fig. 7. Variant 2 state estimation errors for different configurations.

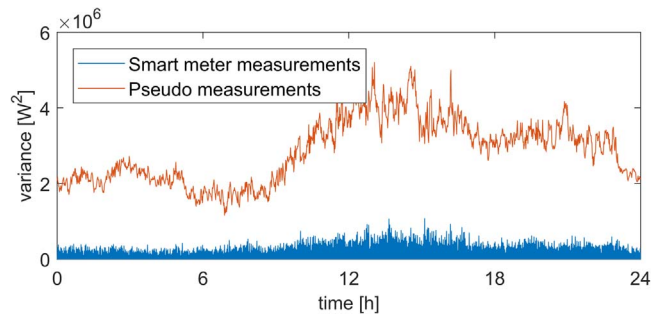


Fig. 8. 2-Minute smart meter measurement variance over time.

various measures on the monitoring accuracy. These measures include the mean error, which is the average error over the three phases, all nodes, the years 2012–2016, and each minute within the 24-h period. The mean maximum error is the maximum error that occurred in one of the nodes (in either of the three phases) during the 24-h period, averaged over the years 2012–2016. The maximum mean error is the maximum of the averaged error over the three phases, all nodes and the years 2012–2016 that occurred within the 24-h period. Finally, the last figure resembles the percentage of time in which the maximum error occurring in one of the nodes (in either of the three phases) was higher than 1%.

The table shows that the base case performs worst in terms of mean error and percentage of time in which the mean error is higher than 1%. However, for the mean maximum error and the maximum mean error, variant 1 (with 15-min time interval synchronised smart meter measurements) performs worse compared to the base case. This can be explained by the fact that, although more measurements are added, these measurement values will contain more extreme values compared to the averaged pseudo measurements. Therefore, these values are causing more extremes in the maximum errors. Obviously, the smaller

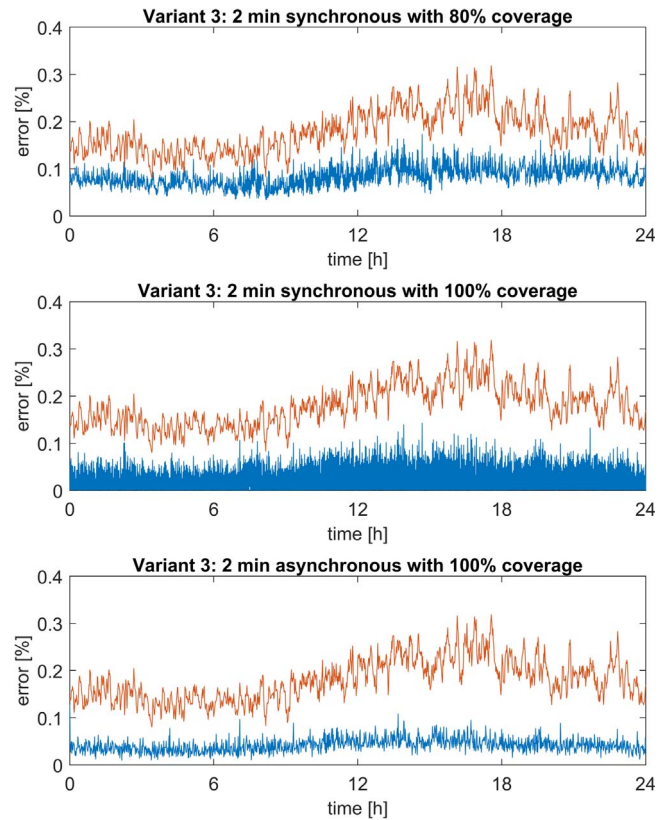


Fig. 9. Variant 3 state estimation errors for different configurations.

the measurement interval, the higher the accuracy for all figures. Mostly, the synchronised measurements perform better in terms of mean error. On the contrary, in most occasions the unsynchronised measurements perform slightly better in terms of mean maximum error, maximum mean error and percentage of time in which the maximum error is higher than 1%. In addition, unsynchronised measuring will also be a more plausible measurement configuration, as there is no requirement for synchronised measurement intervals.

5.2. Legal feasibility and balance of interests

5.2.1. Nature of data

Comparing the above variants in terms of legal feasibility, all variants apart from the base case make use of data that can be classified as personal data. This is mainly due to the fact that the used data is collected from smart meters. Smart meters seem most economic efficient in terms of investment costs. Because at least 80% of the consumers in the EU have to be equipped with smart meters (Annex I(2) Electricity Directive), both the measuring equipment and necessary communication infrastructure should already be in place. Assuming smart meters can be used for monitoring at comparable economic efficiency rates as alternative measuring equipment, smart meter measurements can be produced at little additional costs. Using alternative measurement methods would most likely require installing additional nodes (measuring equipment) in the distribution system, which would be an expensive exercise. Moreover, in order to create the same monitoring accuracy, arguably such nodes would be collecting data on comparable aggregation levels. Consequently, the collected data would still classify as personal data, effectively leading to the same situation as would exist by using smart meter measurements.

5.2.2. Ground for processing

The ground for processing data can be found in the general tasks of the DSO (see Sections 3.3 and 4.2). In the setting of this article,

Table 1
Overall comparison of accuracy for different variants.

	15 min interval			5 min interval			2 min interval			Base case
	80%	100%	100% a	80%	100%	100% a	80%	100%	100% a	
mean error [%]	0,1698	0,1638	0,1699	0,1153	0,0852	0,0990	0,0842	0,0320	0,0416	0,1793
mean max err [%]	1,8833	1,8029	1,8129	1,5005	1,4884	1,3543	1,2588	0,9681	0,8687	1,7727
max mean err [%]	0,3598	0,3495	0,3151	0,2578	0,2549	0,2226	0,1734	0,1425	0,1279	0,3181
time err > 1% [%]	10,4583	10,5417	7,5972	2,2917	1,0417	0,7639	0,6667	0,0278	0,000	12,8611

monitoring accuracy is linked to monitoring local voltage levels. Taking into account the growing amount of decentral RES, strong local voltage variations might occur. In order to deal with these variations in an efficient manner, the DSO should first of all be able to detect the variations. This urges the DSO to monitor at least networks in which many decentral RES are located. As such, the DSO has a ground for processing personal data, up to a minimum to meet the required accuracy for detecting local voltage variations. Arguably, the processed data could also be used for other applications that can increase the security, reliability and efficiency of the DSOs networks. However, each of these applications should be connected to either the general tasks of the DSO or a specific task (see Section 3.3).

5.2.3. Balancing of monitoring accuracy and legal feasibility

When turning to the balancing of interests and minimizing the ‘personal character’ of the data, the following conclusions can be made. Only the 5- and 2-min interval variants, especially with a 100% synchronized coverage, provide a significant improved accuracy compared to the base case, whereas the 15-min time interval variant hardly improves accuracy. Therefore, variant 1 hardly provides additional value in terms of monitoring accuracy compared to the base case and therefore unnecessarily harms household customers’ privacy and is not viable in terms of costs versus benefits. Similarly, variant 3 does provide limited additional value in terms of monitoring accuracy compared to variant 2 and therefore the smaller time intervals of 2 min might be unnecessary, depending on what application the monitoring is used for. As such, it is important to note that other applications might require higher monitoring accuracy and might also justify variant 3. In case the accuracy obtained by variant 2 is sufficient, collecting in shorter timeframes (i.e. 2-min interval) would not be proportionate in terms of customer privacy and an unnecessary burden on communication bandwidth, increasing operational costs. Consequently, if additional accuracy in system monitoring is required (depending on the application it serves), it should be evaluated to what extent the additional benefit of improved accuracy justifies the additional personal data to be collected and additional implementation costs to be made.

For all variants, the processed personal data can be minimized depending on the application it serves. Considering that the state estimation results are used by the DSO for making control actions, short term planning and input for demand side management, the data does not have to be stored for long periods of time, as the related time spans are usually no longer than a day. Afterwards, the data would not be necessary anymore. If the data should be used for other purposes, e.g. network planning analysis, or research purposes, the data can be pseudonymised.

In sum, comparing the cases presented in this work and their ground for data processing, using smart meter data in a 5-min interval with a smart meter coverage of at least 80%, seems to form a reasonable balance in this case study.

6. Conclusion and policy implications

In order to come up with an integrated method to establish distribution system monitoring, this article provides an overview off the different aspects related to establishing distribution network monitoring from both a technical as well as legal perspective. From a technical perspective, it is clear that different variants with different measurement configurations provide highly different monitoring accuracies compared to a base case using only pseudo information as an input. From the overall comparison on the monitoring performance and legal feasibility for the different variants of measurement configurations presented in Section 5, conclusions can be drawn on which variant provides the best balance between technical and legal aspects. For this particular network configuration, households and load profiles for which the simulations have been performed, variant 2 seems to strike a reasonable balance between the monitoring accuracy over time (with a significant improvement compared to the base case) and the impact of corresponding data usage with respect to data protection requirements and the proportionality of the implementation costs in relation to the monitoring accuracy, as discussed in Section 5.

Overall, this article gives insight in what considerations can be made regarding technical and legal issues in establishing distribution network monitoring. For sure several improvements can be made to the various cases presented in this paper. For example, instead of fixed time interval measurements, more dynamic measurements, triggered on certain change rates of the measurement value, can be taken. Further improvements to this work can include a more detailed cost benefit analysis taking into account the benefit of the applications to which the monitoring serves as an input. Nevertheless, it is clear that in order to find a proper balance between accurate network monitoring and privacy of household customers, technicians and lawyers should closely cooperate in assessing the best available options for system monitoring.

Acknowledgement

This research has been funded by The Netherlands Organisation for Scientific Research (NWO) under the DISPATCH project, No. 408-13-056.

References

- Angioni, A., Schlosser, T., Ponci, F., Monti, A., 2016. Impact of pseudo-measurements from new power profiles on state estimation in low-voltage grids. *IEEE Trans. Instrum. Meas.* 65, 70–77. <http://dx.doi.org/10.1109/TIM.2015.2454673>.
- Article 29 Data Protection Working Party, 2011. Opinion 12/2011 on Smart Metering.
- Article 29 Data Protection Working Party, 2007. Opinion 4/2007 on the Concept of Personal Data.
- Bieker, F., Friedewald, M., Hansen, M., Obersteller, H., Rost, M., 2016. A Process for Data Protection Impact Assessment Under the European General Data Protection Regulation. Springer, Cham, pp. 21–37. http://dx.doi.org/10.1007/978-3-319-44760-5_2.
- Blaauwbroek, N., Nguyen, P., Slootweg, H., Nordström, L., 2017. Interfacing solutions for

- power hardware-in-the-loop simulations of distribution feeders for testing monitoring and control applications. *IET Gener. Transm. Distrib.* <http://dx.doi.org/10.1049/iet-gtd.2016.1583>.
- Blaauwbroek, N., Nguyen, P.H., 2015. Conceptual framework and simulation platform for reliable distribution grid monitoring and control with market interaction. In: *Proceedings of the 2015 50th International Universities Power Engineering Conference (UPEC)*, pp. 1–6. <http://dx.doi.org/10.1109/UPEC.2015.7339819>.
- Council of Europe, 2014. Handbook on European Data Protection Law. <http://dx.doi.org/10.2811/53711>.
- Council of Europe, n.d. European Convention on Human Rights.
- de Hert, P., Papakonstantinou, V., 2016. The new general data protection regulation: still a sound system for the protection of individuals? *Comput. Law Secur. Rev.* 32, 179–194. <http://dx.doi.org/10.1016/j.clsr.2016.02.006>.
- Della Giustina, D., Pau, M., Pegoraro, P., Ponci, F., Sulis, S., 2014. Electrical distribution system state estimation: measurement issues and challenges. *IEEE Instrum. Meas. Mag.* 17, 36–42. <http://dx.doi.org/10.1109/MIM.2014.6968929>.
- Diestelmeier, L., Kuiken, D., 2017. Smart electricity systems: access conditions for household customers under EU law. *Eur. Compet. Regul. Law Rev.* 1.
- EDSO for Smart Grids, 2015. European Distribution System Operators for Smart Grids – Response to the European Commission's Public Consultation on a New Energy Market Design.
- European Commission, 2011. Task Force Smart Grids Expert Group 2: Regulatory Recommendations for Data Safety, Data Handling and Data Protection Report. Group.
- European Energy Regulators, 2015. The Future Role of DSOs – A CEER Conclusions Paper.
- European Union, 2016. Regulation 2016/679 of the European parliament and the council of the European Union. *Off. J. Eur. Commun.*
- European Union, 2012a. Charter of Fundamental Rights of the European Union, OJ C 326/02, article 8.
- European Union, 2012b. Consolidated Version of the Treaty on the Functioning of the European Union, OJ C 326/47, article 16(1).
- European Union, 2009. Directive of 2009/72/EC of the European parliament and of the council of 13 July 2009 concerning common rules for the internal market in electricity and repealing directive 2003/54/EC. *Off. J. Eur. Union L211* <http://dx.doi.org/10.1126/science.202.4366.409>. (L 211/55–L 211/93).
- Gungor, V.C., Sahin, D., Kocak, T., Ergut, S., Buccella, C., Cecati, C., Hancke, G.P., 2013. A survey on smart grid potential applications and communication requirements. *IEEE Trans. Ind. Inform.* 9, 28–42. <http://dx.doi.org/10.1109/TII.2012.2218253>.
- IEEE Power and Energy Society, 2015. European Low Voltage Test Feeder [WWW Document]. Digital. http://ewh.ieee.org/soc/pes/dsacom/testfeeders/European_LV_Test_Feeder.zip, (Accessed 3 August 2017).
- McKenna, E., Richardson, I., Thomson, M., 2012. Smart meter data: balancing consumer privacy concerns with legitimate applications. *Energy Policy* 41, 807–814. <http://dx.doi.org/10.1016/j.enpol.2011.11.049>.
- Milaj, J., Mifsud Bonnici, J.P., 2016. Privacy Issues in the Use of Smart Meters—Law Enforcement Use of Smart Meter Data. Springer International Publishing, Cham, Switzerland, pp. 179–196. http://dx.doi.org/10.1007/978-3-319-28077-6_12.
- Pau, M., Pegoraro, P.A., Sulis, S., 2013. Efficient branch-current-based distribution system state estimation including synchronized measurements. *IEEE Trans. Instrum. Meas.* 62, 2419–2429. <http://dx.doi.org/10.1109/TIM.2013.2272397>.
- Pecan Street Inc., 2016. Pecan Street Inc. | Energy Research [WWW Document]. <http://www.pecanstreet.org/energy/>, (Accessed 3 August 2017).
- Pérez-Arriaga, L.J., 2013. *Regulation of the Power Sector*, Springer London, Power Systems. Springer, London.
- Singh, R., Pal, B.C., Vinter, R.B., 2009. Measurement placement in distribution system state estimation. *IEEE Trans. Power Syst.* 24, 668–675. <http://dx.doi.org/10.1109/TPWRS.2009.2016457>.
- Torbaghan, S.S., Blaauwbroek, N., Nguyen, P., Gibescu, M., 2016. Local market framework for exploiting flexibility from the end users. In: *Proceedings of the 2016 13th International Conference on the European Energy Market (EEM)*, pp. 1–6. <http://dx.doi.org/10.1109/EEM.2016.7521304>.
- Verbong, G.P.J., Verkade, N., Verhees, B., Huijben, J.C.C.M., Höffken, J.I., 2016. Smart Business for Smart Users: A Social Agenda for Developing Smart Grids. Springer International Publishing, Cham, Switzerland, pp. 27–42. http://dx.doi.org/10.1007/978-3-319-28077-6_3.
- Wang, H., Schulz, N.N., 2004. A revised branch current-based distribution system state estimation algorithm and meter placement impact. *IEEE Trans. Power Syst.* 19, 207–213. <http://dx.doi.org/10.1109/TPWRS.2003.821426>.